Genetic Engineering Approaches for Enhanced Production of Biodiesel Fuel from Microalgae

P. G. Roessler, L. M. Brown, T. G. Dunahay, D. A. Heacox, E. E. Jarvis, J. C. Schneider, S. G. Talbot, and K. G. Zeiler

National Renewable Energy Laboratory, 1617 Cole Blvd., Golden, CO 80401

Microscopic algae ("microalgae") have great potential as a future source of biological lipids for use in biodiesel fuel production. We are investigating the biochemistry and molecular biology of lipid biosynthesis in these organisms, and are developing genetic transformation systems that should allow us to genetically engineer microalgal strains for enhanced biodiesel production capabilities. As a result of these studies, we have isolated a full-length gene for the key lipid biosynthetic enzyme acetyl-CoA carboxylase. In addition, we have isolated portions of the genes that encode nitrate reductase and acetolactate synthase as a first step in the development of homologous selectable markers for microalgal transformation.

Diminishing global reserves of petroleum have prompted a search for alternative, renewable liquid fuels. One type of alternative fuel receiving consideration is derived from biological lipids. This fuel, referred to as "biodiesel," is produced by a simple transesterification process that converts glycerolipids into methyl or ethyl esters of fatty acids, along with glycerol as a byproduct. A small but burgeoning biodiesel industry is emerging, using oilseeds and waste animal fats as the primary lipid sources. Facilities with a total biodiesel production capacity of 32 million gallons (125 million liters) per year are currently operating in Europe, where high taxation of petroleum-derived fuels has stimulated the market for biodiesel (*1*). There are plans to increase the annual capacity of European facilities to more than 200 million gallons (800 million liters) in the next few years. On the other hand, the low cost of petroleum-derived diesel fuel in the United States has resulted in very limited commercial activity involving biodiesel in this country.

Environmental concerns are playing a major role in promoting the commercialization of biodiesel (*2*). Biodiesel is a cleaner burning fuel than conventional diesel. Because of the naturally low concentration of sulfur in biodiesel, the production of sulfur oxides during combustion is greatly reduced.

Emissions of carbon monoxide and particulates are also lower for biodiesel. Consequently, the use of biodiesel in urban areas of the United States that have poor air quality may help these areas to attain compliance with the Clean Air Act Amendments of 1990. Furthermore, unlike petroleum-derived diesel, biodiesel exhibits low toxicity and is biodegradable. This latter property has focused attention on the advantages of using lipid-based oils and hydraulic fluids in environmentally sensitive areas. Another benefit of biodiesel is that it is produced photosynthetically, using carbon dioxide as the primary feedstock. This "greenhouse gas" is thereby recycled, resulting in a lower rate of atmospheric accumulation.

The small current market for biodiesel fuel can be satisfied by existing oilseed production capacities. As the market for biodiesel increases, however, additional sources of biological lipids will be needed (*3*). These new sources will have to be produced in large quantities, and must not compete for resources (e.g., land and water) that are currently being used for conventional agriculture. There should also be minimal competition for the resources that will be needed for the production of lignocellulosic energy crops for use as a feedstock for ethanol fuel production.

Microscopic eukaryotic algae ("microalgae") are promising candidate organisms that may be able to fulfill the need for an additional lipid source. Certain strains of microalgae are able to produce up to 60% of their total cellular mass as lipid (*4*). Of equal importance is the fact that many microalgae grow extremely well in saline water that is unsuitable for crop irrigation. In addition, many microalgal strains are able to tolerate widely fluctuating temperatures and high light intensities. Consequently, the desert regions of the southwestern United States, which have large stores of saline groundwater, could be used for microalgal mass culture. These areas cannot be used for conventional agriculture, meaning that competing land uses would be minimal. Furthermore, because carbon dioxide is required in large quantities as a nutrient for microalgal mass culture, it is worth noting that coal-burning electric power plants are located in close proximity to some of the areas deemed suitable for microalgal mass culture. The carbon dioxide present in the flue gas produced by these facilities would therefore be available for utilization by microalgae for biodiesel production (*5*).

Massive accumulation of lipids in microalgae typically occurs only when the cells are environmentally stressed, and under these conditions the growth rates of cultures are greatly reduced. Nutrient limitation is one type of stress that often leads to lipid accumulation coupled to reduced growth rates (*4*). It would be desirable to manipulate a microalgal strain so that it could maintain a relatively high growth rate and produce greater quantities of lipids at the same time. One possible way to achieve this goal is through genetic engineering. This paper describes some of the research being carried out at the National Renewable Energy Laboratory (NREL) in an effort to develop microalgal strains that have enhanced biodiesel production capabilities. Current research efforts are focused primarily on the physiology, biochemistry, and molecular biology of oil-producing microalgae.

Biochemical Studies

The main goal of our biochemical research is to better understand the enzymatic pathways involved in microalgal lipid biosynthesis. A detailed knowledge of these pathways is necessary prior to any attempts to genetically engineer microalgae for altered lipid production capabilities. We are interested in identifying enzymes that influence the overall rate of lipid synthesis and the quality of the lipids produced. Identification of rate-controlling enzymes may eventually allow us to manipulate the activity of these enzymes in a manner that leads to enhanced lipid accumulation. Likewise, characterization of the key steps involved in fatty acid modification may allow us to control the nature of the lipids produced, which has implications regarding the quality of the final fuel product.

Fatty Acid Modification. We recently began a project designed to examine the biochemical processes that control the chain length and level of unsaturation of fatty acids in microalgae. Chain length and double bond number affect fuel quality factors such as oxidative stability, ignitability, and combustion performance (*6*). For instance, the thermal efficiency decreases with increasing carbon chain length (*7*). Likewise, although a single carbon-carbon double bond in the acyl chain helps to reduce the freezing point and increase the thermal efficiency of the fuel (*7*), greater numbers of double bonds impair the combustion performance (*6*). The presence of more than two double bonds also decreases the oxidative stability of biodiesel fuel. An ideal diesel fuel consists of medium to long unbranched hydrocarbon chains containing a single double bond (*7*).

Storage triacylglycerols (TAGs) from many microalgae are composed primarily of C_{14} to C_{18} fatty acids that are saturated or monounsaturated (*4*). The methyl or ethyl esters of these fatty acids would therefore function very well as a diesel fuel. Conversely, membrane lipids in most microalgae are rich in elongated and polyunsaturated fatty acids, and therefore are less desirable as a biodiesel feedstock. Because these less desirable fatty acids comprise from 20 to 80 mol % of the total (depending on the species and growth conditions), it would be advantageous to be able to manipulate the chain length and unsaturation level of membrane lipid fatty acids in order to improve fuel quality. In this respect, it is worth noting that mutants of the higher plant *Arabidopsis thaliana* have been isolated that are defective in certain fatty acid desaturation steps, and yet no obvious physiological effects of these mutations are exhibited under normal growth conditions. One such mutation decreases the proportion of linolenic acid (18:3)[1] in leaves from 49% to 19%, with a concomitant increase in the level of linoleic acid (18:2) (*8*).

The unicellar marine alga *Nannochloropsis* is an excellent organism in which to study microalgal fatty acid elongation and desaturation. One strain of

---

[1]Standardized fatty acid nomenclature is utilized throughout the text, where the number to the left of the colon represents the acyl chain length and the number to the right of the colon represents the number of double bonds in the chain.

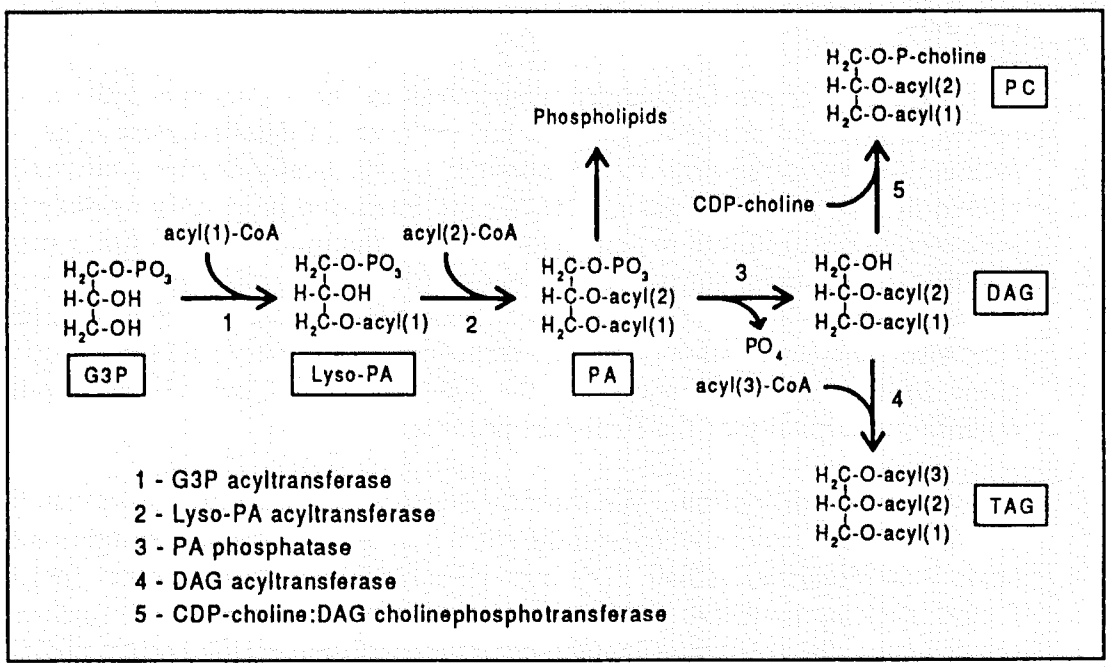

Figure 1. Pathways of glycerolipid biosynthesis.

Nannochloropsis has been shown to accumulate up to 55% of its cell mass as lipid under nitrogen-deficient conditions (*9*). Furthermore, light intensity affects the ratio of storage TAGs to photosynthetic membrane lipids in *Nannochloropsis*, thereby altering the ratio of medium/long chain, saturated/monounsaturated fatty acids to very long chain, polyunsaturated fatty acids (*10*). This model system should allow us to examine various biochemical aspects of fatty acid modification.

The synthesis of TAGs and many polar membrane lipids occurs within the microsomal membranes via the pathway shown in Figure 1. Two acyltransferases are required to produce phosphatidic acid (PA) from glycerol-3-P (G3P) and acyl-coenzyme A (acyl-CoA) molecules. This PA can be used directly as a substrate for polar lipid synthesis, or can be dephosphorylated to diacylglycerol (DAG), which is the substrate for the synthesis of TAG, phosphatidylcholine (PC), and galactolipids. It should be noted that PC can undergo an acyl exchange reaction at the *sn*-2 position with acyl-CoA; this reaction is catalyzed by lyso-PC acyltransferase. This latter point is important, in that fatty acids esterified to PC in this position are subject to desaturation by a PC-specific desaturase in higher plants.

The acyltransferases involved in microsomal lipid synthesis in higher plants have been shown to exhibit pronounced preferences for specific acyl-CoA molecules (e.g., *11*, *12*). The acyl specificities exhibited by the lyso-PA and DAG acyltransferases are even more pronounced when the glycerolipid substrates contain certain fatty acids. These acyl specificities, along with specificities for particular glycerolipid substrates exhibited by enzymes involved in polar lipid biosynthesis, are believed to play a major role in determining the final acyl compositions of the various lipid classes. Thus, our initial studies involving the biochemistry of lipid biosynthesis in *Nannochloropsis* were designed to provide information on acyltransferase substrate specificities.

Microsomal membrane preparations were made from mid-log phase *Nannochloropsis* cells by a procedure that includes grinding the cells in the presence of glass beads, low-speed centrifugation to remove cellular debris, and ultracentrifugation at 100,000 x g to pellet the membranes. The membranes were incubated with [^{14}C]palmitoyl-CoA (16:0-CoA) or [^{14}C]oleoyl-CoA (18:1-CoA) in the presence of G3P for specific periods of time, after which the lipids were extracted with organic solvents and separated by one-dimensional thin layer chromatography (TLC) to separate neutral lipids or two-dimensional TLC to separate polar lipids. Areas of the TLC plates containing radioactivity (as determined by autoradiography) were scraped into vials for scintillation counting. These experiments indicated that 16:0-CoA is incorporated most readily into PA, while 18:1-CoA is incorporated most readily into PC (see Figure 2). The labeling of PC is probably due to acyl exchange with previously synthesized PC, rather than to *de novo* synthesis of PC, since CDP-choline (a substrate needed for PC synthesis) was not included in the assay mixture. In higher plants, 18:1 acyl groups incorporated into PC are usually subjected to further desaturation. PC may serve as a site for desaturation in *Nannochloropsis* as well, since the bulk of cellular unsaturated C_{18} fatty acids are found in PC. Preliminary results also suggest that 16:0-CoA is incorporated more strongly into PA than 18:1-CoA, leading to proportionally more label in DAG and TAG. The apparent difference in incorporation of 16:0-CoA and 18:1-CoA into PA implies that acyl chain specificity is exhibited for one or both of the acyltransferases that form this lipid. Although lyso-PA is present in very small quantities in membranes, a small amount of labeled compound that comigrates with this lipid intermediate was observed on TLC plates. This lipid was labeled to about the same extent regardless of whether 16:0-CoA or 18:1-CoA was provided (data not shown), suggesting that the enzyme that catalyzes the synthesis of lyso-PA (G3P acyltransferase) does not have a strong preference for one of these acyl chains over the other. If true, then the specificity for formation of PA must be exercised at the second step, lyso-PA acyltransferase. Indirect evidence has also suggested that the last enzyme of TAG biosynthesis, DAG acyltransferase, does not have a strong preference for particular acyl-CoA molecules.

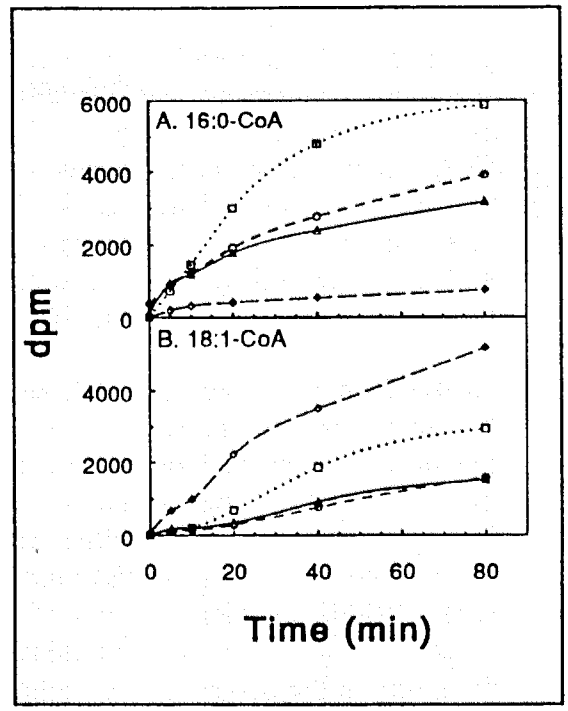

Figure 2. Incorporation of acyl-CoAs into lipids in *Nannochloropsis sp.* Symbols: (○) DAG; (△) TAG; (□) PA: and (◇) PC.

Our preliminary conclusions are that the enzyme that acylates the second position of the glycerol backbone, lyso-PA acyltransferase, has a high substrate specificity, whereas G3P acyltransferase and DAG acyltransferase are less discriminating. Additionally, lyso-PC acyltransferase prefers 18:1-CoA over 16:0-CoA. Confirmation of these conclusions will require substrate competition studies in which mixtures of acyl-CoAs (including acyl chain lengths greater than 18 carbons) are provided to the enzymes. *In vivo* pulse-chase experiments using various labeled substrates are also being carried out to analyze the pathways of lipid biosynthesis and fatty acid modification.

Acetyl-CoA Carboxylase. We have also carried out research to identify enzymes that play a role in controlling the overall rate of lipid biosynthesis in microalgae. For these studies, we have utilized the diatom *Cyclotella cryptica*. The cell walls of diatoms contain substantial amounts of silica, and thus this nutrient is required in large quantities. When silica becomes depleted from the medium, growth of *C. cryptica* cultures slows and lipids (primarily storage TAGs) accumulate rapidly (*13*). Within the first 4 hours of silicon deficiency, there is a doubling in the percentage of newly photoassimilated carbon that is partitioned into lipids, and a reduction in the percentage of carbon that is partitioned into storage carbohydrate. Using this model system, we have obtained evidence that acetyl-CoA carboxylase (ACCase) may play a role in this lipid accumulation process, in that the activity of this enzyme doubles after 4 hours of silicon deficiency and quadruples after 15 hours (*14*). This increase in activity can be blocked by the addition of protein synthesis inhibitors, suggesting control at the level of gene expression.

ACCase is a biotin-containing enzyme that catalyzes the ATP-dependent carboxylation of acetyl-CoA to form malonyl-CoA. Two partial reactions are involved in this process: 1) carboxylation of the enzyme-bound biotin molecule, and 2) transfer of the carboxyl group to acetyl-CoA. In the bacterium *Escherichia coli*, these reactions are catalyzed by four distinct, separable protein components: 1) biotin carboxylase, 2) biotin carboxyl carrier protein, and 3) two subunits of carboxyltransferase. In eukaryotic organisms, these entities are located on a single, multifunctional polypeptide typically having a molecular mass exceeding 200 kilodaltons (kDa) (*15*). The functional ACCase enzyme in eukaryotes is composed of multimers of this large polypeptide.

ACCase is believed to catalyze the rate-limiting step of fatty acid biosynthesis in animals (*16, 17*). The rate of fatty acid synthesis in higher plants also appears to be dependent on the activity of ACCase (*18, 19*). In addition, studies with *E. coli* have suggested that ACCase may catalyze the rate-limiting step of phospholipid biosynthesis (*20*). It is also worth noting that the expression of the ACCase gene in *E. coli* is correlated to the cellular growth rate (*21*).

In light of the regulatory significance of ACCase in lipid biosynthesis, we sought to purify this enzyme from *C. cryptica* and to determine some of its biochemical properties. The enzyme was purified 600-fold to near homogeneity by a combination of ammonium sulfate precipitation, gel filtration chromatography, and monomeric avidin affinity chromatography (*22*). The molecular mass of the native enzyme was determined to be between 700 and 800 kDa, while the subunit mass

was determined to be approximately 200 kDa; this suggested that the functional enzyme was composed of four identical, biotin-containing subunits. Various cellular metabolites were tested for their ability to modulate the activity of the purified ACCase (*23*). ACCase activity was inhibited somewhat in the presence of the reaction products (malonyl-CoA, PO_4, and ADP), and inhibited more strongly in the presence of palmitic acid and palmitoyl-CoA. Pyruvate stimulated enzyme activity slightly (~50%) at high (1 mM) concentrations. The physiological significance of these effects is not known at this time.

Although we have determined some of the structural and catalytic properties of ACCase, characterization of various molecular aspects of the enzyme (e.g., primary structure, active site geometry, etc.) requires the isolation and cloning of the ACCase gene. A cloned copy of the gene is also required for research concerning the regulation of ACCase gene expression. Our efforts to clone this gene are discussed in the following section.

Genetic Engineering

As discussed earlier, genetic engineering is one approach for producing microalgal strains with enhanced biodiesel production capabilities. Two primary requirements must be fulfilled before a strain can be successfully engineered in this manner: 1) an expressible, recombinant gene that affects lipid metabolism must be available; and 2) methods must be developed to stably incorporate this cloned gene into a host cell in a manner that allows expression of the gene.

ACCase Gene Cloning. Because ACCase plays an important role in controlling the rate of fatty acid biosynthesis in many organisms, this enzyme has been receiving an increasing amount of attention as a target for manipulation via genetic engineering for the purpose of increasing the lipid production capabilities of various organisms. Therefore, we attempted to clone this gene from *C. cryptica*. When this project first started, the ACCase gene had only been isolated from rat (*24*) and chicken (*25*). The gene has now also been isolated from yeast (*26*) and all of the subunits of *E. coli* ACCase have been cloned (*27-30*). There have not been any previous reports of a full-length ACCase gene from eukaryotic algae or higher plants.

ACCase is relatively easy to purify from *C. cryptica*. Therefore, we chose to take a polymerase chain reaction (PCR)-based approach to clone the ACCase gene from this species (*31*). PCR primers were designed from partial amino acid sequences that were determined for various CNBr-generated fragments of the purified enzyme. These PCR primers were used to amplify a 146-base pair (bp) fragment of the ACCase gene, using DNA isolated from *C. cryptica* as a template. This PCR product was subsequently used to identify a full-length clone of the gene in a genomic *C. cryptica* library.

The complete DNA sequence of the ACCase gene from *C. cryptica* has now been determined. This sequence has been deposited in the GenBank genetic data base (accession number L20784). In addition, the sequences of several regions of the ACCase gene transcript (i.e., messenger RNA) have been determined, which has

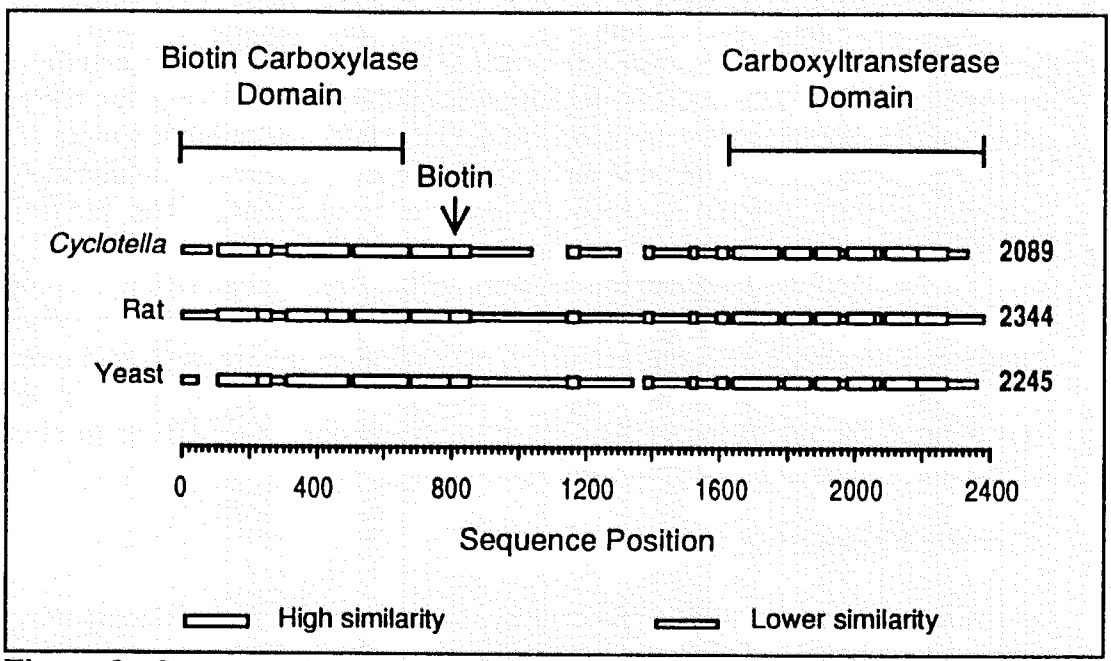

Figure 3. Sequence similarities among the ACCase enzymes from *C. cryptica*, rat, and yeast. Based on the MACAW program of Schuler et al. (32). The biotin attachment site is indicated by an arrow.

allowed identification of the probable translation initiation codon, along with the precise localization of two introns in the gene. A 73-bp intron is present near the region of the gene that encodes the biotin-binding site of the enzyme, and a 447-bp intron is located near the start of the coding region of the gene. Upon removal of these introns, a coding sequence 6.3 kilobases long is obtained, which encodes a protein that contains 2089 amino acids and has a molecular mass of approximately 230 kDa. *C. cryptica* ACCase is therefore comparable in size to the multifunctional ACCases that have been isolated from various plants, animals, and yeast. The large size is necessary because of the three functional domains that are responsible for the different partial reactions involved in malonyl-CoA synthesis: biotin carboxylase, biotin carboxyl carrier protein, and carboxyltransferase.

Comparison of the deduced amino acid sequences of ACCase from *C. cryptica*, yeast, and rat and indicated strong similarities in the biotin carboxylase and carboxyltransferase domains (~50% identity), but less similarity in the middle region of the enzyme (~30% identity). This middle region, which includes the biotin carboxyl carrier protein domain, probably functions as a "swinging arm" to transfer the carboxylated biotin from the biotin carboxylase domain to the carboxyltransferase domain. As such, a fair amount of variability in the amino acid sequence of this region could be tolerated, thus explaining the comparatively lower sequence conservation. Regions with statistically significant similarity are shown in Figure 3, which is the graphical output generated by the MACAW sequence alignment computer program developed by Schuler et al. (*32*).

The amino-terminal sequence of the predicted ACCase polypeptide, shown below using the standard one-letter amino acid code, has characteristics of a "signal sequence" (*33*). Note the two positively charged arginine residues within the first few amino acids of the polypeptide, followed by a hydrophobic region:

MALRRGLYAAAATAILVTASVTAFAPQHSTFTPQSLSAAP. . .

Signal sequences direct polypeptides into the endoplasmic reticulum (ER) in eukaryotic cells. In diatoms, nuclear-encoded proteins destined for chloroplasts must be transported through the ER (*34*), which is consistent with the fact that diatom chloroplasts are completely enclosed by ER membranes (*35*). Because fatty acid biosynthesis occurs primarily in the chloroplasts of higher plants (*36*), we have assumed that ACCase is located in the chloroplasts of diatoms, and therefore a signal sequence would be necessary for chloroplast targeting. Alternatively, it is possible that the ACCase gene that we have cloned codes for an ER-localized ACCase that is responsible for producing the malonyl-CoA utilized for elongation of fatty acids from C_{16} or C_{18} to C_{20} and C_{22}. Because ACCase must pass through the ER before it can enter the chloroplast, it is also possible that a single ACCase isoform is functional in both of these cellular compartments. Further experimentation will be necessary to examine this possibility. The location of the cleavage site at the end of the putative ACCase signal sequence is not clear; attempts to determine the amino-terminal sequence of the mature enzyme have not been successful.

We have begun to manipulate the ACCase gene in order to produce forms that can be expressed in a variety of organisms. For example, to prevent possible problems with inefficient intron excision in heterologous hosts, we have removed the two introns by replacing fragments of the genomic DNA with corresponding fragments of cDNA. We have also performed site-specific mutagenesis of the gene to add certain restriction sites and eliminate others, thereby facilitating the construction of expression vectors. In addition, we are altering the 5' end of the coding region in order to change the sequence at the amino terminus of the expressed polypeptide; these changes will be useful for directing the enzyme into particular subcellular compartments of the host organisms. For example, to direct *C. cryptica* ACCase into the chloroplasts of higher plants, it will be necessary to add a gene fragment that encodes a chloroplast transit peptide. On the other hand, for host organisms in which fatty acid synthesis is cytoplasmic (e.g., yeast), it will be necessary to remove the signal sequence-encoding DNA from the gene.

Development of Genetic Transformation Systems for Microalgae. The isolation of genes that affect lipid accumulation in microalgae is the first step toward genetically engineering microalgal strains for enhanced biodiesel production. It is then necessary to incorporate these genes into the host cell in a stable manner that allows efficient expression of the gene. In recent years, the process of genetic transformation, defined here as the introduction and expression of exogenous genes in a host organism, has become routine for many bacterial, fungal, animal, and plant species (*37, 38*). However, stable transformation systems are not currently available

for most microalgae. A major focus of our research is the development of genetic transformation systems for microalgae that have potential with respect to biodiesel fuel production. The two basic steps in this process, development of methods to introduce DNA into algal cells and detection of expression of the exogenous gene, are discussed in detail below.

Introduction of Foreign DNA. For both algae and higher plants, the cell wall is a significant barrier to the introduction of DNA or other macromolecules into the cell. For many plant species, the cell wall can be removed enzymatically to form protoplasts. Protoplasts can be induced to take up DNA through the cell membrane by a variety of methods, including electroporation (*39*) or treatment with polyethylene glycol (*40*). Previous work in our laboratory involving the transient expression of firefly luciferase has suggested that DNA can be introduced into algal cells by standard techniques if algal protoplasts can be formed (*41*). However, the production of protoplasts from most strains of microalgae is not straightforward. Currently, the only microalgal species for which an efficient transformation system exists is the green alga *Chlamydomonas reinhardtii.* Transformation of *C. reinhardtii* was facilitated by the availability of wall-less cells, produced either by genetic mutation (*42*) or through the use of autolysin, a species-specific cell wall-degrading enzyme produced by gametes during mating (*43*). Kindle and coworkers (*44*) found that DNA could be introduced into these wall-less cells simply by agitating the cells in the presence of glass beads. Unfortunately, wall-less mutants, or methods to generate wall-less cells, are not available for most microalgal species. The composition of microalgal cell walls is highly variable between species, and even between different strains of the same species. Furthermore, the cell walls often contain components, such as sporopollenin, that are resistant to degradation by the enzyme mixtures commonly used to degrade plant cell walls. To avoid the difficult work involved in developing protocols to generate viable protoplasts from each algal species of interest, a simple technique for introduction of DNA into intact (walled) algal cells is preferable.

Recently, several methods have been developed to introduce DNA into cells without prior removal of the cell wall. The technique of microprojectile bombardment, or biolistics, uses pressurized helium gas to propel DNA-coated gold particles into target cells or tissues, and has been used successfully to transform a large variety of cell types, including *C. reinhardtii* (*44, 45*). The glass bead agitation method described above has also been used successfully to introduce DNA into walled cells of yeast (*46*) and walled cells of *C. reinhardtii* (*43*), although the transformation frequency was low. Recently, two groups found that DNA uptake in higher plant cells could be facilitated by agitation of the cells with silicon carbide fibers (SiC) (*47, 48*). We have adapted the SiC protocol for the transformation of walled cells of *C. reinhardtii* (*49*) and found that the transformation efficiency was similar to that achieved using glass beads or particle bombardment. However, the SiC method may have advantages over the other protocols. For example, the cell survival rate was much higher using SiC as compared to glass beads, suggesting a more "gentle" mechanism. Furthermore, expensive equipment is not required for the SiC method, unlike the particle bombardment technique. We are currently

working to adapt the SiC protocol for transformation of microalgal cells that have greater potential for fuel production.

Selectable Marker Genes. Even with the best DNA introduction technology, the fraction of cells in which gene entry and stable expression are achieved is very low. Therefore, marker genes are required to facilitate detection of these rare transformation events. These selectable markers also aid in the stabilization of introduced sequences, since cells that lose the marker gene can be selected against. In *C. reinhardtii*, successful transformation has required the use of homologous genes (*44, 50*). Although heterologous genes (e.g., bacterial marker genes) have been expressed at low levels in *C. reinhardtii* (*51*), they have not been effective as selectable markers in algae. This is probably due to poor expression of non-algal genes, possibly because of differences in algal promoter signals, codon bias, or DNA methylation. The DNA of *C. reinhardtii* has an unusually high guanosine plus cytosine (GC) content, which is reflected in an unusual bias toward these bases in codon usage (e.g., *52, 53*). Furthermore, a recent report suggests that DNA methylation may affect the expression of foreign genes in *C. reinhardtii* (*54*).

We are working to develop homologous genes as selectable markers for the transformation of certain species of microalgae that have potential utility for fuel production. The analysis of DNA from these organisms indicates that some microalgae, particularly the green algae, exhibit high GC contents and high degrees of cytosine methylation (*55*). The DNA of one of the target species, *Monoraphidium minutum*, has a GC content of 71%. This indicates that its codon usage may be biased toward G and C bases, as in *C. reinhardtii*, suggesting that heterologous selectable marker genes with a lower, more typical GC content may not be effective in *M. minutum*. By using homologous genes, problems with promoter specificity and codon bias will not inhibit gene expression, allowing successful selection of transformed cells. We are exploring the development of three different homologous selectable marker systems for the transformation of oil-producing microalgae.

Nitrate reductase. The nitrate reductase (NR) gene has been used in other laboratories as a transformation marker. In *C. reinhardtii*, NR-deficient mutant cells have been transformed successfully with the wild-type *C. reinhardtii* NR gene using the glass bead and SiC procedures described above. Progress has been made in the generation of mutants and in the cloning of the NR gene from *M. minutum* in hopes of developing a homologous transformation system similar to that used for *C. reinhardtii*.

Algal mutants lacking functional NR can be selected based on their resistance to chlorate (*56, 57*). Cells having a functional NR enzyme will take up chlorate (a nitrate analog) and reduce it to chlorite, which is toxic to the cells. The NR-deficient mutants do not reduce the chlorate and so are not subject to the cytotoxic effects of chlorite. Consequently, cells can be subjected to a positive selection regime by growing target organisms in the presence of chlorate and picking out resistant colonies. Using this protocol, we have isolated several putative NR-deficient mutants of *M. minutum*. These candidates grow in the

presence of chlorate and are unable to utilize nitrate as a nitrogen source. Biochemical analyses are under way to confirm that this phenotype is in fact due to mutations in the NR structural gene.

The next step in the development of this particular transformation system for *M. minutum* is the isolation of the functional NR gene from the wild-type strain. This gene will be used to complement the NR-deficient strains; putative transformants will be selected by their ability to use nitrate as the sole nitrogen source. A genomic library of *M. minutum* DNA was constructed in a lambda vector. To obtain a homologous probe with which to screen the library, PCR primers were designed based on conserved sequences in the coding regions of NR genes from three other green algae: *C. reinhardtii* (Lefebvre, P., Univ. of Minn., personal comm., 1992), *Volvox carterii* (*58*), and *Chlorella vulgaris* (*59*). PCR amplification of total DNA from *M. minutum* produced a 750-bp fragment that was cloned and partially sequenced, confirming that the clone represents a portion of the NR gene. This NR gene fragment is currently being used to screen the *M. minutum* genomic library in order to obtain the full-length NR gene.

Orotidine-5'-phosphate decarboxylase. The orotidine-5'-phosphate decarboxylase (OPD) gene codes for an essential enzyme in the pyrimidine biosynthesis pathway. OPD-deficient mutants grow only if provided with an alternate source of pyrimidines, such as uracil, or if they have been transformed with a functional copy of an OPD gene. There is also a positive selection for OPD mutants; OPD converts the compound 5-fluoroorotic acid (FOA) into a toxic compound, killing wild-type cells, whereas OPD mutants can grow on FOA-containing media (*60*). A spectrophotometric assay can be used to measure OPD activity in cell extracts (*61*). The OPD gene has been cloned and sequenced from a number of organisms (e.g., *62-65*), and has been used successfully in several transformation systems (e.g., *66-68*).

Use of the FOA mutant selection method has allowed us to obtain several putative OPD mutants from ultraviolet-mutagenized *M. minutum* cells. These cells require uracil for growth and exhibit a low reversion rate, thus providing a powerful selection for cells that become transformed with the wild-type OPD gene. Work is in progress to isolate the wild-type OPD gene from *M. minutum*. Two approaches are being taken. The first relies on the PCR technique, using degenerate primers designed from regions of the gene that appear to be conserved between the OPD genes of various species. The second approach is to clone the gene directly by functional complementation of *E. coli* or *Saccharomyces cerevisiae* OPD mutants.

Herbicide resistance. Using herbicide resistance as a selectable marker has the advantage of traditional antibiotic selection in that resistance genes can be used as dominant selectable markers. In other words, wild-type organisms would be able to be transformed directly without the need to generate mutant strains, which is a process that can be particularly problematic in diploid cells such as diatoms. This method requires either 1) cloning and *in vitro* site-specific mutagenesis of the wild-type herbicide-sensitive gene to an herbicide-resistant form, or 2) cloning of the

gene from a mutant strain that has become resistant to the herbicide. Herbicide resistance genes that have been developed in plant systems include 5-enol-pyruvylshikimate-3-phosphate, which is responsible for glyphosate resistance, and acetolactate synthase (ALS), which is involved in resistance to sulfonylurea and imidazolinone herbicides (*69*). These systems have potential as selectable markers for the transformation of microalgae as well. In fact, several species in the NREL Culture Collection, including *M. minutum* and *C. cryptica*, have been shown to be sensitive to such herbicides (*70*). Efforts are under way to clone and mutate the microalgal genes involved in resistance to these herbicides. For the ALS gene, a PCR-based approach similar to that described above for the NR gene was used to amplify a ~500-bp gene fragment from the DNA of *M. minutum;* this fragment exhibits strong sequence similarity to the ALS genes of other organisms. This partial gene will be used as a probe to isolate the full-length gene from a *M. minutum* genomic library.

Conclusions

Microalgae are a potential source of biological lipids for use in large-scale biodiesel production. We have made substantial progress in identifying some of the factors that are important in determining the quantity and quality of lipids produced by microalgae. This work has led to the cloning of ACCase, which is believed to play an important role in controlling the rate of fatty acid biosynthesis. Progress is also being made with respect to the development of methods to genetically engineer oil-producing microalgae, with the eventual goal of enhanced biodiesel production.

Acknowledgments

We are grateful to J. Bleibaum and G. Thompson of Calgene, Inc. for performing amino acid sequence analysis of ACCase. A significant portion of the ACCase gene cloning research was carried out in collaboration with J. Ohlrogge at Michigan State University; this support is gratefully acknowledged. This work was supported in part by the Biofuels Systems Division of the United States Department of Energy, the National Science Foundation, and the United States Department of Agriculture-Binational Agricultural Research and Development Fund.

Literature Cited

1. *Biodiesel Alert* **1993,** *7(1).*
2. Shay, E. G. *Biomass and Bioenergy* **1993,** *4*, 227-242.
3. Brown, L. M. In *Proceedings of the First Biomass Conference of the Americas, Vol II* (NREL/CP-200-5768); National Renewable Energy Laboratory, Golden, Colorado, 1993; pp. 902-909.
4. Roessler, P. G. *J. Phycol.* **1990,** *26*, 393-399.
5. Chelf, P.; Brown, L. M.; Wyman, C. E. *Biomass and Bioenergy* **1993,** *4*, 175-183.
6. Harrington, K. J. *Biomass* **1986,** *9*, 1-17.

7. Klopfenstein, W. E.; Walker, H. S. *J. Amer. Oil Chem. Soc.* **1983,** *60*, 1596-1598.
8. Browse, J.; McCourt, P.; Somerville, C. *Plant Physiol.* **1986,** *81*, 859-864.
9. Suen, Y.; Hubbard, J. S.; Holzer, G.; Tornabene, T. G. *J. Phycol.* **1987,** *23*, 289-296.
10. Sukenik, A.; Carmeli, Y.; Berner, T. *J. Phycol.* **1989,** *25*, 686-692.
11. Oo, K. -C.; Huang, A. H. C. *Plant Physiol.* **1989,** *91*, 1288-1295.
12. Griffiths, G.; Stobart, A. K.; Stymne, S. *Biochem. J.* **1985,** *230*, 379-388.
13. Roessler, P. G. *Arch. Biochem. Biophys.* **1988,** *267*, 521-528.
14. Roessler, P. G. *J. Phycol.* **1988,** *24*, 394-400.
15. Samols, D.; Thornton, C. G.; Murtif, V. L.; Kumar, G. K.; Haase, F. C.; Wood, H. G. *J. Biol. Chem.* **1988,** *263*, 6461-6464.
16. Kim, K. -H.; Lopez-Casillas, F.; and Bai, D. -H. *FASEB J.* **1989,** *3*, 2250-2256.
17. Lane, M. D.; Moss, J.; Polakis, S. E. In *Current Topics in Cellular Recognition*; Horecker, B. L.; Stadtman, E. R., Eds.; Academic Press, New York, NY, 1974, Vol. 8; pp 139-195.
18. Post-Beittenmiller, D.; Jaworski, J.G.; Ohlrogge, J.B. *J. Biol. Chem.* **1991,** *266*, 1858-1865.
19. Post-Beittenmiller, D.; Roughan, G.; Ohlrogge, J. B. *Plant Physiol.* **1992,** *100*, 923-930.
20. Jackowski, S.; Cronan Jr., J. E.; Rock, C. O. In *Biochemistry of Lipids, Lipoproteins and Membranes*; Vance, D. E.; Vance, J., Eds.; Elsevier Science Publishers B.V., Amsterdam, 1991; pp 43-85.
21. Li, S. -J.; Cronan Jr., J. E. *J. Bacteriol.* **1993,** *175*, 332-340.
22. Roessler, P. G. *Plant Physiol.* **1990,** *92*, 73-78.
23. Roessler, P. G.; Ohlrogge, J. B. In *The Biochemistry, Structure, and Utilization of Plant Lipids*; Harwood, J.; Quinn, P., Eds.; Portland Press, London, 1990; pp 108-110.
24. Lopez-Casillas, F.; Bai, D. -H.; Luo, X.; Kong, I. -S.; Hermodson, M. A.; Kim, K. -H. *Proc. Natl. Acad. Sci. U.S.A.* **1988,** *85*, 5784-5788.
25. Takai, T.; Yokoyama, C.; Wada, K.; Tanabe, T. *J. Biol. Chem.* **1988,** *263*, 2651-2657.
26. Al-Feel, W.; Chirala, S. S.; Wakil, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89*, 4534-4538.
27. Li, S. -J.; Cronan Jr., J. E. *J. Biol. Chem.* **1992,** *267*, 855-863.
28. Li, S. -J.; Cronan Jr., J. E. *J. Biol. Chem.* **1992,** *267*, 16841-16847.
29. Kondo, H.; Shiratsuchi, K.; Yoshimoto, T.; Masuda, T.; Kitazono, A.; Tsuru, D.; Anai, M.; Sekiguchi, M.; Tanabe, T. *Proc. Natl. Acad. Sci. U.S.A.* **1991,** *88*, 9730-9733.
30. Alix, J.-H. *DNA* **1989,** *8*, 779-789.
31. Roessler, P. G.; Ohlrogge, J. B. *J. Biol. Chem.* **1993,** *268*, 19254-19259.
32. Schuler, G. D.; Altschul, S. F.; Lipman, D. J. *Proteins Struct. Funct. Genet.* **1991,** *9*, 180-190.
33. von Heijne, G. *J. Membrane Biol.* **1990,** *115*, 195-201.

34. Bhaya, D.; Grossman, A. *Mol. Gen. Genet.* **1991,** *229*, 400-404.
35. Gibbs, S. P. *J. Cell. Sci.* **1979,** *35*, 253-266.
36. Harwood, J. L. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **1988,** *39*, 101-138.
37. Fincham, J. R. S. *Microbiol. Reviews* **1989**, *53*, 148-170.
38. Potrykus, I. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **1991**, *42*, 205-225.
39. Fromm, M.; Taylor, L. P.; Walbot, V. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5824-5828.
40. Krens, F. A.; Molendijk, L.; Wullems, G. J.; Schilperoort, R. A. *Nature* **1982**, *296*, 72-74.
41. Jarvis, E. E.; Brown, L. M. *Curr. Genet.* **1991**, *19*, 317-321.
42. Davis, D. R.; Plaskitt, A. *Genet. Res.* **1971**, *17*, 33-43.
43. Kindle, K. L. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 1228-1232.
44. Kindle, K. L.; Schnell, R. A.; Fernández, E.; Lefebvre, P. A. *J. Cell Biol.* **1989**, *109*, 2589-2601.
45. Klein, T. M.; Wolf, E. D.; Wu, R.; Sanford, J. C. *Nature* **1987**, *327*, 70-73.
46. Costanzo, M. C.; Fox, T. D. *Genetics* **1988**, *120*, 667-670.
47. Asano, Y.; Otsuki, Y.; Ugaki, M. *Plant Science* **1991**, *79*, 247-252.
48. Kaeppler, H. F.; Somers, D. A.; Rines, H. W.; Cockburn, A. F. *Theo. Appl. Genet.* **1992**, *84*, 560-566.
49. Dunahay, T. G. *Biotechniques* **1993,** *15*, 452-460.
50. Debuchy, R.; Purton, S.; Rochaix, J. -D. *EMBO J.* **1989**, *8*, 2803-2809.
51. Hall, L. M.; Taylor, K. B.; Jones, D. D. *Gene* **1993**, *124*, 75-81.
52. Goldschmidt-Clermont, M.; Rahire, M. *J. Mol. Biol.* **1986**, *191*, 421-432.
53. Silflow, C. D.; Youngblom, J. *Ann. N.Y. Acad. Sci.* **1986**, *466*, 18-30.
54. Blankenship, J. E.; Kindle, K. L. *Mol. Cell. Biol.* **1992**, *12*, 5268-5279.
55. Jarvis, E. E.; Dunahay, T. G.; Brown, L. M. *J. Phycol.* **1992**, *28*, 356-362.
56. Cove, D. J. *Mol. Gen. Genet.* **1976**, *146*, 147-159.
57. Toby, A. L.; Kemp, C. L. *J. Phycol.* **1977**, *13*, 368-372.
58. Gruber, H.; Goetinck, S. D.; Kirk, D. L., Schmitt, R. *Gene* **1992**, *120*, 75-83.
59. Cannons, A. C.; Iida, N.; Solomonson, L. P. *Biochem. J.* **1991**, *278*, 203-209.
60. Boeke, J. D.; LaCroute, F.; Fink, G. R. *Mol. Gen. Genet.* **1984**, *197*, 345-346.
61. Donovan, W. P.; Kushner, S. R. *J. Bacteriol.* **1983**, *156*, 620-624.
62. Rose, M.; Grisafi, P.; Botstein, D. *Gene* **1984**, *29*, 113-124.
63. Newbury, S. F.; Glazebrook, J. A.; Radford, A. *Gene* **1986**, *43*, 51-58.
64. Ohmstede, C.; Langdon, S. D., Chae, C.; Jones, M. E. *J. Biol. Chem.* **1986**, *261*, 4276-4282.
65. Turnbough, C. L.; Kerr, K. H.; Funderburg, W. R.; Donahue, J. P.; Powell, F. E. *J. Biol. Chem.* **1987**, *262*, 10239-10245.
66. Buxton, F. P.; Radford, A. *Mol. Gen. Genet.* **1983**, *190*, 403-405.

67. Boy-Marcotte, E.; Vilaine, F.; Camonis, J.; Jacquet, M. *Mol. Gen. Genet.* **1984**, *193*, 406-413.
68. van Hartingsveldt, W.; Mattern, I. E.; van Zeijl, C. M. J.; Pouwels, P. H.; van den Hondel, C. A. M. *Mol. Gen. Genet.* **1987**, *206*, 71-75.
69. Botterman, J.; Leemans, J. *Trends Genet.* **1988**, *4*, 219-222.
70. Galloway, R. E. *J. Phycol.* **1990**, *26*, 752-760.